

Fig 2a: Pass Rate vs Model Size (Scaling Behavior)

